Purposeful Statistical Investigations

Lorraine Day
University of Notre Dame Australia
<Lorraine.Day@nd.edu.au>

Lorraine Day provides us with a great range of statistical investigations using various resources such as maths300 and TinkerPlots. Each of the investigations link mathematics to students' lives and provide engaging and meaningful contexts for mathematical inquiry.

The *Australian Curriculum: Mathematics* (Australian Curriculum Assessment and Reporting Authority (ACARA), 2013) states: "[The students] represent, summarise and interpret data and undertake purposeful investigations involving the collection and interpretation of data" (p. 2). The use of the word "purposeful" is important here. What does it mean to be purposeful? In creating tasks and investigations the purpose is to introduce particular mathematical ideas and provide opportunities for students to make sense of them. Rather than focussing on procedures such as algorithms, rules and techniques, it is important to build into tasks the need for students to use mathematical ideas in ways that will allow them to see the usefulness of the mathematics both within a particular task and to solve other associated problems. By considering purpose as well as context, the teacher is able to make stronger connections between school mathematics and the real world by considering why and how mathematical ideas are powerful (Ainley, 2012). Consideration should be given to whether the activity has the potential to promote the required reflective thought to enable students to construct understanding of the mathematical relationships. The following are a collection of purposeful statistical investigations that teachers might like to try in their classrooms.

Using students' personal data to form a 'picture': What does our class like?

Students are more interested in data which have meaning for them or are about them. Building a 'picture' of the likes, dislikes and interests of students in a class allows students to work with data that they find meaningful. The purpose of the first activity described is to collect data, organise it into categories and create displays using a variety of graphing techniques to allow interpretation and comparison of data displays.

The activity begins with the collection of class data, in this case favourite ice-cream flavours, using a tally (see Figure 1). Konold and Kazak (2008) talk about the importance of linking chance to data at appropriate opportunities, so a discussion about the probability of a random student chosen from the class liking chocolate ice cream, could be considered at this point.

Favourite Ice Cream			
Chocolate	|||	3	$\frac{3}{27}$
Mint choc chip	⧘⧘	5	$\frac{5}{27}$
Cookies 'n cream	⧘⧘ |||	8	$\frac{8}{27}$
Vanilla	⧘⧘	5	$\frac{5}{27}$
Cookie dough	⧘⧘		6	$\frac{6}{27}$
		27	

Figure 1. Tally of favourite ice-cream flavours.

Figure 2. Strip graph of favourite ice cream flavours.

Discussion about how to display this data may result in students talking about picture graphs, column graphs and pie charts. A strip graph is a simple way to look at the data. It can be made on a strip of cardboard, as shown in Figure 2. Conversations with children about why each ice cream needs to be represented by an equal division on the strip graph can be quite illuminating. It is important that students can see that the length of the strip graph represents the whole (class). This allows the fractions and percentages looked at later in the activity to be related back to the whole.

The strip graph can then be made into a circle by joining the two ends and taping them together. This way the students can see the connection between the linear strip graph and the circle which forms the basis of the pie chart. By placing the circular graph on a piece of poster paper, students are able to trace around the inside and mark off where each flavour appears on the circumference (see Figures 3 and 4).
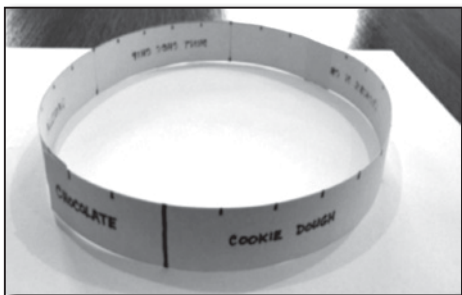

Figure 3. Trace around the inside of the circle.

Figure 4. Draw in the divisions and label.

A 100-bead string can then be placed around the circle for the percentages to be estimated for each sector of the graph (see Figures 5 and 6). This provides the link between the fractions and percentages, which many students fail to see. The link to probability may again be made by comparing the pie chart to a spinner.
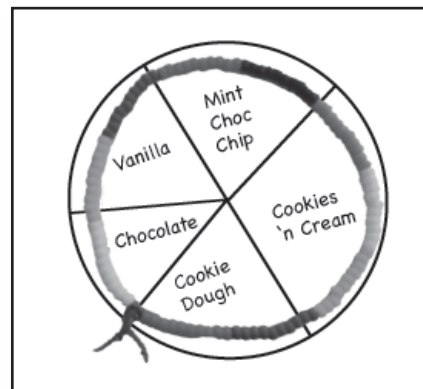

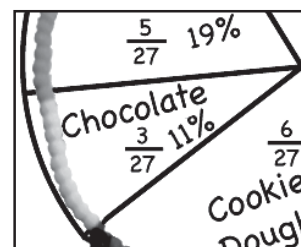Figure 5. Place the 100-bead string.

Figure 6. Read the percentage from the bead string.

By cutting up the strip graph, a column graph can also be made. The clear connection between the pie graph, column graph and strip graph as three different representations of the same data can be discussed.

Once the students have had this experience they are able to collect several sets of data about the class, the types of which the students could decide. For example, what are the students' favourite TV shows, music genres, movies and fast foods? By combining all of this information, a picture of what the class likes may be drawn up. The data can be recorded by completing a survey and posting in bins (see Figure 7).

This is a quick and easy method of collecting quite a large amount of data in a no-fuss manner. The students complete a survey, cut up the results and place in the appropriate bin. Then each group may be given one of the bins to collate, represent and analyse the data. When all the groups put the collected data together, the result is a 'picture' of what the entire class likes.
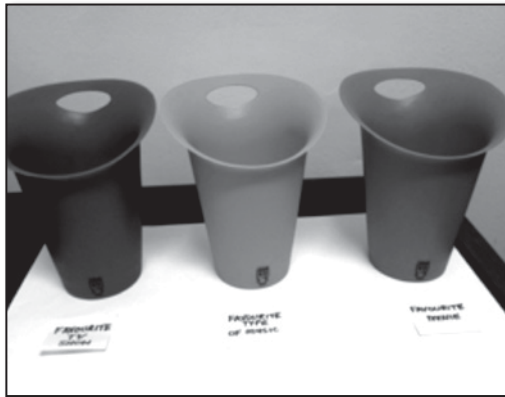

Figure 7. Data collection bins.

This activity begins simply and without the need for any difficult calculations; therefore it is accessible to all students, while developing an understanding of how pie charts and column graphs may be developed from a simple strip graph. This activity was inspired by Lesson 11 *'This Goes With That'* from *maths300* (Education Services Australia, 2010).

Another method of collating class data is by using a software program, such as *TinkerPlots* (Konold & Miller, 2004), where the data about each student is recorded on a data card and can be retrieved at any time.

TinkerPlots was purposefully designed for students in the middle years of schooling, but is not beyond students in the primary school setting. It allows students to investigate different forms of graphical representation in order to decide which displays the data in the best possible way for the data type. Figure 8 shows two representations of the favourite TV data that students may create and compare to determine which graph best represents the data. By experimenting with the different data displays and evaluating the effectiveness of each, the students can create a 'picture' of what their class likes.

Comparing class data to population data: What does our class look like?

Using *TinkerPlots* to build up a data card on each student in the class allows many other data and relationships to be investigated. For example, data such as hand span, arm span, height, eye colour, hair colour, and shoe size can all be added to the data card. Students are then able to build up a picture of what their class looks like, and perhaps draw inferences about what the typical Year 4, 5 or 6 student might look like. They could compare their data to data on the *CensusAtSchool* website provided by the Australian Bureau of Statistics (n.d.). This would provide an opportunity to investigate the differences between samples and populations as well as introducing random sampling, sample size and the reason for the census.
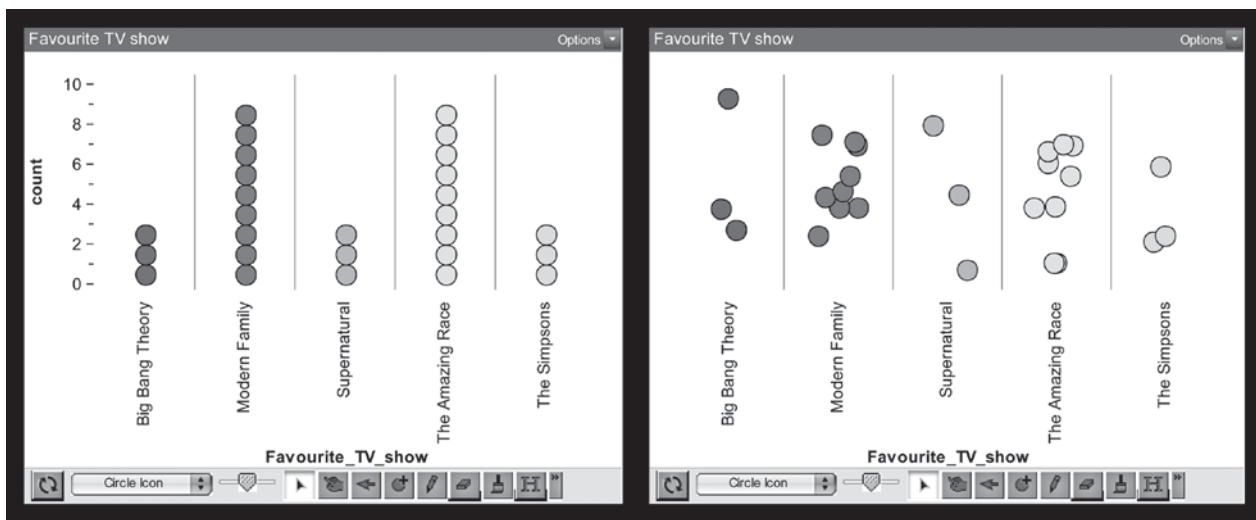

Figure 8. Students can experiment to determine which display best represents the data.

Watson et al. (2011) devote an entire chapter of *Digging Into Australian Data with TinkerPlots* to investigating data about school students using *CensusAtSchool* data (see Figure 9) with many investigations that may be used with student data. These lessons could be modified to use data collected at the class or school level as well as utilising the *CensusAtSchool* database. These lessons cover many of the big ideas of statistics including: hypothesising, using measures of central tendency (mean, median and mode) and spread to interpret data, collecting random samples and the importance of sample size, relationships, and causation. Different representations such as stacked dot plots, pie charts, hat plots, box plots, scatter plots and lines of best fit are introduced as students need to interpret, analyse and make sense of their data.

Figure 9. Data about school students from CensusAtSchool.

Using data to reason mathematically: Identifying the mystery attributes

One type of activity that is really useful for emphasising the Reasoning proficiency from the *Australian Curriculum: Mathematics* (ACARA, 2013) is identifying the mystery attribute. In these lessons, students are provided with graphs using some of the attributes on the corresponding data cards (see Figure 10). They are asked to identify which attributes the graphs are using, the reason they have come to this conclusion, and then duplicate the graph using the appropriate attributes.

Promoting understanding: Using sharing to find the mean

So often students are provided with an algorithm to find the mean of a set of numbers (add the scores and divide by the number of scores) without really appreciating what the mean is representing. A favourite activity to find the mean of a set of numbers is a silent sharing game. Each student in a group is given a number of blocks and students are asked to share the blocks between them so they all have the same number.
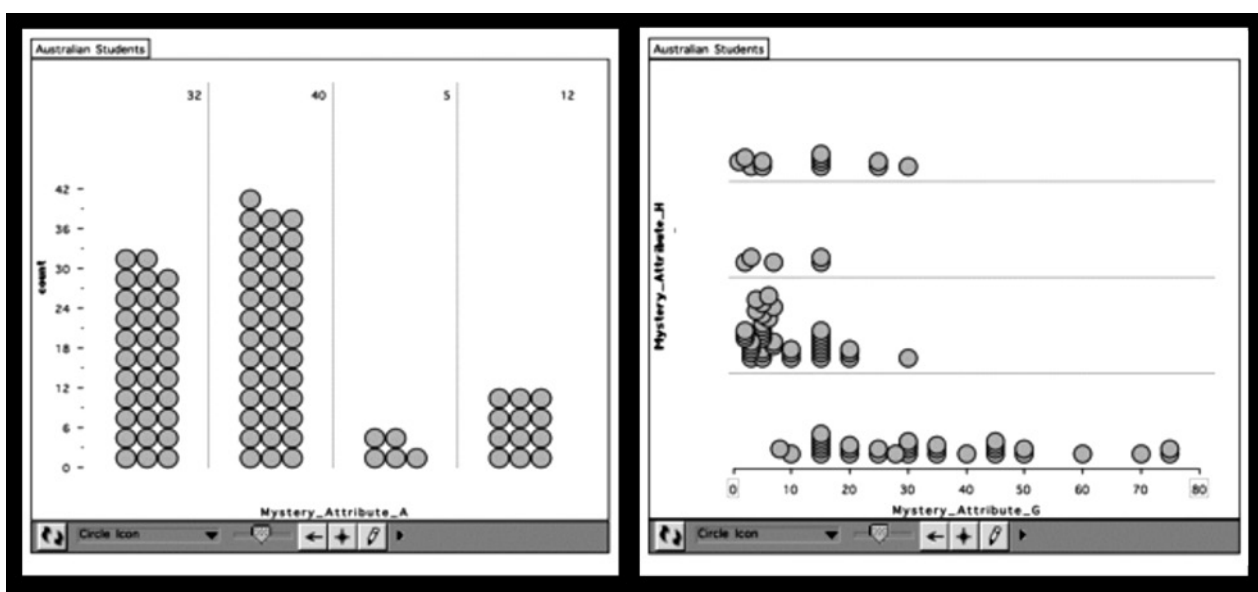

Figure 10. Examples of identifying the mystery attributes (Data file for Lesson 3.2 (Watson et al., 2011).

The rules of the game are that it takes place in silence and that students may give blocks to others but may not take blocks from others (see Figure 11). The rest of the class watches on in silence and the mental calculations going on across the whole class are obvious as they watch the active group. It is fascinating to see how students arrange the blocks so that it is easy for other group members to see how many they have. Asking students to reflect on the process and explain what would happen if one person was given a lot more blocks or no blocks demonstrates a deep understanding of the idea of the mean.
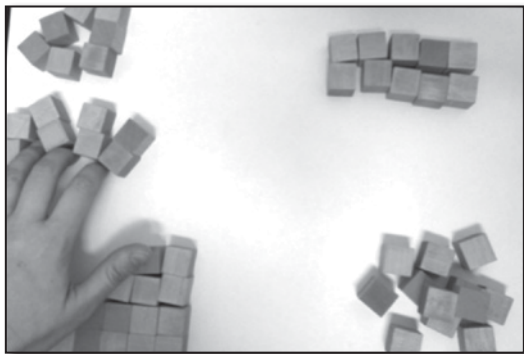

Figure 11. Sharing to find the mean.

Incorporating the proficiency strands: What is typical?

Another type of activity that emphasises the Understanding and Reasoning proficiencies, while describing and interpreting different data sets within a context, is exploring a data set to see what is 'typical'. In using this activity with a group of students, they particularly enjoyed investigating what were the attributes of a typical Melbourne Cup winner (Watson et al., 2011). By interrogating the data, they found that the typical winner was a bay stallion, either four or five years old, which carried around 51–52 kg with a starting price about 9 or 10 to 1 (see Figure 12). Depending on whether they used a range of values or the median for starting prices and weights carried, different groups were able to justify their choices of the horses they thought were the most typical. An incredible amount of quality mathematical discussion as groups justified their selections to each other was evident.
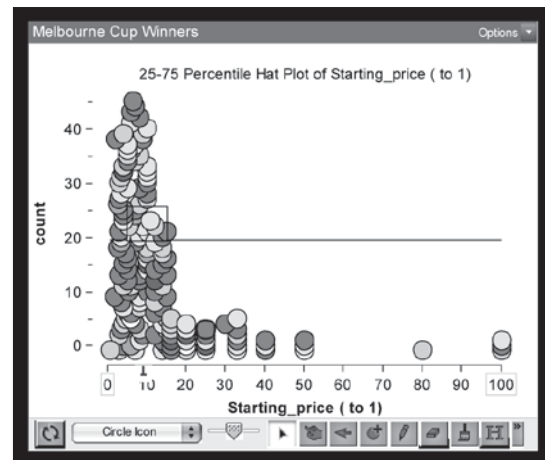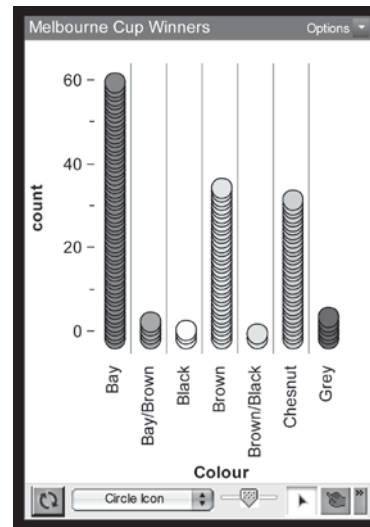

Figure 12. Examples of graphs used to find the typical Melbourne Cup winner (data file for lesson 4.2 in Watson et al., 2011).

Creating mathematical models that relate to students' interests

To support the development of informal inference, Konold and Kazak (2008) suggest that the study of statistics and probability should be linked so that, rather than calculations, there is a focus on the shape of probability distributions through collection of data and simulations. A task that does this and that students seem to enjoy is *Dice Footy* from the *maths300* website (Education Services Australia, 2010). In this lesson, students model a game of AFL football with dice, where one die represents the number of goals in a quarter and the other the number of behinds. As with all *maths300* lessons, it is important for students to play the game and collect their own data (see Figure 13) before moving to a computer simulation to collect even more data. Students play several games

of *Dice Footy* and collect the class data in an appropriate manner. A stem and leaf plot works well for this purpose, as a general shape of the distribution emerging can be seen. Students are encouraged to 'tell the story' of the distribution.
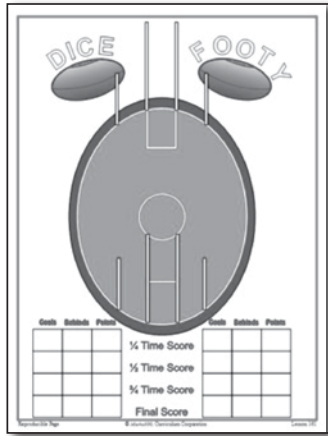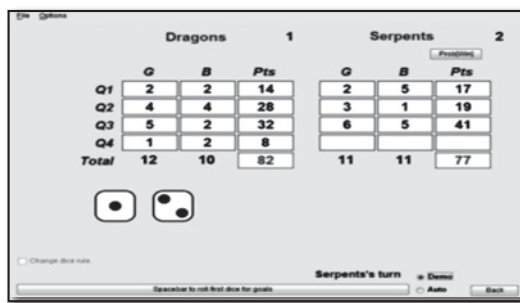

Figure 13. Dice Footy board.

Figure 14. Computer simulation of Dice Footy (Education Services Australia, 2010).

The computer simulation duplicates what the students have done with the dice to demonstrate that the computer plays the game in exactly the same manner. A timely teaching idea is to pause the game when the first team has completed their game and the second team is about to play the final quarter (see Figure 14) and ask the students what the second team must roll in order to win. To do this, students need to demonstrate fluency, understanding and reasoning.

The computer simulation allows many games to be played and the average numbers of goals, behinds and points calculated. These values provide results remarkably close to real games of AFL football, which allows AFL statistics to be investigated and compared to the mathematical model. It also allows the exploration of expected values with older students. Another aspect of the simulation is that the point or goal distribution may be graphed. By beginning with a small number of games and slowly building up to a sample of 100 000 games, students are able to 'tell the story' of the graphs as more data are collected and they can see the shape of the distribution being formed (see Figure 15). There are many other ways to use the computer simulation, such as investigating the possibilities of draws, or looking at the possible range of scores, expected values and the limitations of this particular model.
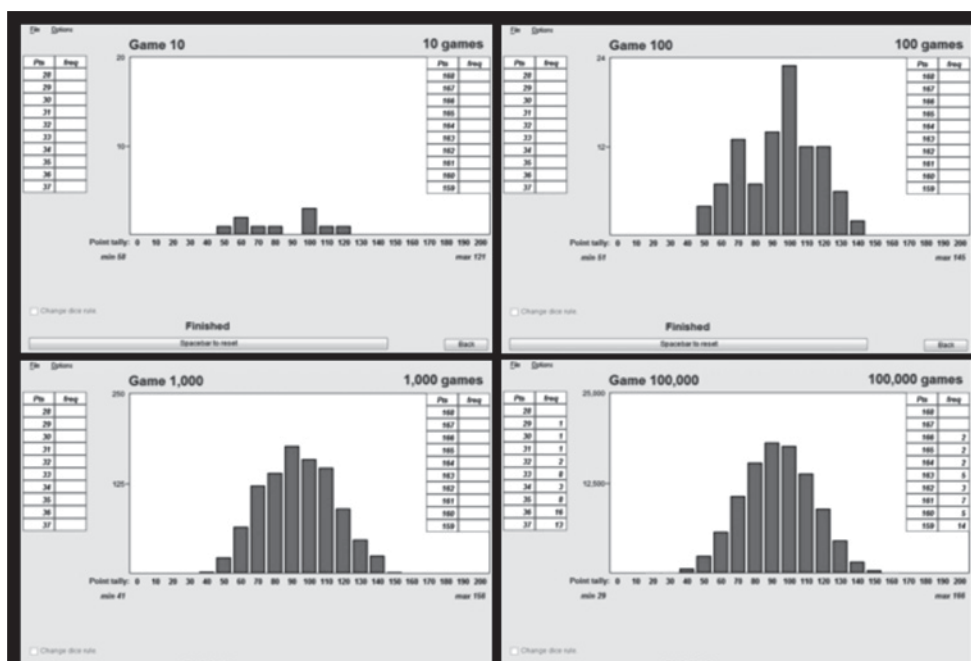

Figure 15. Points distributions for different numbers of games of Dice Footy.

Conclusion

In the primary school classroom, the Statistics and Probability strand can sometimes be trivialised into a few semi-engaging opportunities to collect a bit of data, much of which is of no real interest to students nor serves to show the capacity of mathematics to make sense and order out of students' lives. Rather, this strand should be an opportunity to involve students in rich investigative tasks that relate to real life, other learning areas and important social issues. T*he Australian Curriculum: Mathematics* (ACARA, 2013) states:

> Statistics and Probability initially develop in parallel and the curriculum then progressively builds the links between them. Students recognise and analyse data and draw inferences. They represent, summarise and interpret data and undertake purposeful investigations involving the collection and interpretation of data. They assess likelihood and assign probabilities using experimental and theoretical approaches. They develop an increasingly sophisticated ability to critically evaluate chance and data concepts and make reasoned judgments and decisions, as well as building skills to critically evaluate statistical information and develop intuitions about data. (p. 2)

The tasks represented here are really just the 'tip of the iceberg' of the richness available for teachers in this strand. They serve to illustrate how purposeful investigations not only satisfy the proficiency strands but provide engaging and meaningful contexts for mathematical inquiry.

References

Ainley, J. (2012). *Developing purposeful mathematical thinking: A curious tale of apple trees.* PNA, 6(3), 85–103. Retrieved from http://hdl.handle.net/10481/19524

Australian Bureau of Statistics [ABS]. (n.d.) *CensusAtSchool.* Retrieved from http://www.abs.gov.au/censusatschool

Australian Curriculum Assessment and Reporting Authority [ACARA]. (2013). *Australian curriculum: Mathematics.* Retrieved from http://www.australiancurriculum.edu.au/Mathematics/Content-structure

Education Services Australia [ESA]. (2010). *maths300.* Retrieved from http://www.maths300.esa.edu

Konold, C., & Kazak, S. (2008) Reconnecting data and chance. *Technology Innovations in Statistics Education, 2*(1), 1–37.

Konold, C. & Miller, C. (2004). *TinkerPlots: Dynamic data exploration.* Emeryville, CA: Key Curriculum Press.

Watson, J., Beswick, K., Brown, N., Callingham, R., Muir, T., & Wright, S. (2011). *Digging into Australian data with TinkerPlots: Data analysis for middle school students.* Sandown Village, Vic.: Objective Learning Materials.